

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

日 本 国 特 許 庁

21.07.00

PATENT OFFICE
JAPANESE GOVERNMENT

REC'D 12 SEP 2000

WIPO PCT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年 7月23日

出 願 番 号

Application Number:

平成11年特許願第209817号

出 願 人

Applicant(s):

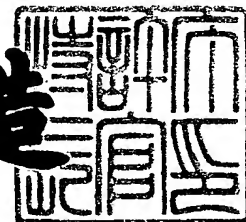
株式会社ヘリックス研究所

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2000年 8月25日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2000-3066564

【書類名】 特許願
 【整理番号】 H1-108
 【提出日】 平成11年 7月23日
 【あて先】 特許庁長官 殿
 【国際特許分類】 C12N 15/00
 【発明者】

【住所又は居所】 神奈川県藤沢市辻堂新町 1-2-7-105

【氏名】 太田 紀夫

【発明者】

【住所又は居所】 茨城県稲敷郡阿見町大室 511-12

【氏名】 磯貝 隆夫

【発明者】

【住所又は居所】 東京都板橋区氷川町 27-3-403

【氏名】 西川 哲夫

【発明者】

【住所又は居所】 千葉県木更津市矢那 4508-19-201

【氏名】 河合 弓利

【特許出願人】

【識別番号】 597059742

【氏名又は名称】 株式会社ヘリックス研究所

【代理人】

【識別番号】 100102978

【弁理士】

【氏名又は名称】 清水 初志

【選任した代理人】

【識別番号】 100108774

【弁理士】

【氏名又は名称】 橋本 一憲

【手数料の表示】

【予納台帳番号】 041092

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9716404

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 全長 cDNA、およびタンパク質

【特許請求の範囲】

【請求項 1】 配列番号：2、配列番号：4、配列番号：6、および配列番号：8 から選択される、いずれかの配列番号として記載されたアミノ酸配列からなるタンパク質。

【請求項 2】 請求項 1 に記載のいずれかのタンパク質をコードする DNA。

【請求項 3】 配列番号：1、配列番号：3、配列番号：5、および配列番号：7 から選択される、いずれかの配列番号として記載された塩基配列のコード領域からなる、請求項 2 に記載の DNA。

【請求項 4】 請求項 2 または 3 に記載の DNA のいずれかが挿入されたベクター。

【請求項 5】 請求項 2 または 3 に記載の DNA のいずれかを発現可能に保持する形質転換体。

【請求項 6】 請求項 5 に記載の形質転換体を培養し、発現産物を回収する工程を含む、請求項 1 に記載のタンパク質のいずれかを製造する方法。

【請求項 7】 請求項 2 または 3 に記載の DNA のいずれか、またはその相補鎖にハイブリダイズする DNA であって、少なくとも 15 ヌクレオチドの鎖長を持つ DNA。

【請求項 8】 請求項 7 に記載の DNA からなる、請求項 1 に記載のタンパク質をコードするヒト全長 cDNA 合成用プライマー。

【請求項 9】 請求項 7 に記載の DNA からなる、請求項 1 に記載のタンパク質をコードする遺伝子の検出用プローブ。

【請求項 10】 請求項 2 または 3 に記載のいずれかの DNA もしくはその一部に対するアンチセンス DNA。

【請求項 11】 次の工程を含む、請求項 1 に記載のタンパク質をコードする全長 cDNA の合成方法。

a) cDNA ライブラリーを鋳型として請求項 8 に記載のプライマーを起点とする相補鎖合成反応を行い、

b) 合成産物を回収する

【請求項 12】 cDNA ライブラリーが、オリゴキャップ法によって合成されたものである請求項 11 に記載の合成方法。

【請求項 13】 相補鎖の合成を PCR 法によって行う請求項 11 に記載の合成方法。

【請求項 14】 請求項 1 に記載のいずれかのタンパク質に対する抗体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ヒトに由来するタンパク質をコードする全長 cDNA、この cDNA によってコードされるタンパク質、並びにそれらの製造および用途に関する。

【0002】

【従来の技術】

現在、世界的なレベルで様々な生物のゲノム配列の解明とその解析が進められている。既に 10 種類を越える原核微生物、下等真核生物の出芽酵母、多細胞性真核生物である線虫で、その全ゲノム配列が決定された。3,000,000,000 塩基対といわれるヒトのゲノムについては、現在、世界的な協力体制のもとでその解析が進められており、2002～2003 年頃までには、その全構造が明らかにされようとしている。ゲノム配列を明らかにする目的は、複雑な生命現象をその設計図であるゲノム情報を解読し、全ての遺伝子の機能や制御、あるいは遺伝子間、タンパク質間、細胞間さらには個体間における相互作用のネットワークとして生物を理解するところにある。種々の生物種のゲノム情報から生命現象を解明していくことは、単に学術分野における研究課題として重要であるのみならず、そこで得られる研究成果をいかに産業上の応用へと発展させていくかという点で、その社会的な意義も大きい。

ところが単にゲノム配列を決定しただけでは、全ての遺伝子の機能を明らかにできるわけではない。例えば酵母では、ゲノム配列から推定された約 6,000 の遺伝子の約半数しか、その機能を推定できなかった。一方、ヒトには約 100,000 種類の遺伝子が存在するといわれる。そこで、ゲノム配列から明らかにされてくる

膨大な量の新しい遺伝子の機能を、迅速かつ効率的に解明していくための「ハイスループット遺伝子機能解析システム」の確立が、強く望まれている。

【0003】

真核生物のゲノム配列では、多くの場合、一つの遺伝子がイントロンによって複数のエキソンに分断されている。そのため、ゲノム配列情報だけからそこにコードされるタンパク質の構造を正確に予測するには、多くの問題がある。一方、イントロンが除かれたmRNAから作製されるcDNAでは、タンパク質のアミノ酸配列の情報が一つの連続した配列情報として得られるため、容易にその一次構造を明らかにすることが可能である。ヒトのcDNAの研究では、これまでに1,000,000を超えるEST (Expression Sequence Tags) データがパブリックドメインに公開されており、それらはヒトの全遺伝子の80%以上をカバーしているものと推定されている。

これらの情報は、ヒト遺伝子構造の解明やゲノム配列におけるエキソン領域の予測、あるいはその発現プロファイルの推定など、様々な角度から利用されている。ところが、これらのヒトEST情報の多くはcDNAの3' 末端側近傍に集中しているため、特にmRNAの5' 末端近傍の情報が極端に不足している状況にある。また、これらのヒトcDNAの中でコードされているタンパク質の配列が予測されているmRNAは約7,000種類程度であり、更にそのうち全長cDNAクローンとして取得されているものはわずか5,500種類程度に過ぎないのが現状である。ESTとして登録されているものを含めても、全長クローンとしてこれまでに取得されているヒトcDNAは、ヒト全遺伝子のわずかに10%~15%程度であると推定されている。

【0004】

完全長cDNAでは、その5' 末端配列からゲノム配列上でのmRNA転写開始点が特定できる上、その配列の中に含まれるmRNAの安定性や翻訳段階での発現制御に関わる因子の解析が可能である。また、翻訳開始点であるatgを5' 側に含むことから、正しいフレームでタンパク質への翻訳を行うことができる。したがって、適当な遺伝子発現系を適用することで、そのcDNAがコードするタンパク質を大量に生産したり、タンパク質を発現させてその生物学的活性を解析することも可能になる。このように、完全長cDNAの解析からはゲノム配列解析を相補する重要な情

報が得られる。また、発現可能な全長cDNAクローンは、その遺伝子の機能の実証的な解析や産業分野での応用への展開において、その重要性はきわめて高い。

【0005】

全長cDNAを合成する方法は公知である。たとえばオリゴキャップ法 [K. Maruyama and S. Sugano, Gene, 138: 171-174 (1994); Y. Suzuki et al., Gene, 200: 149-156 (1997)] によれば、原理的には全長cDNAに富むライブラリーを合成することができる。合成したcDNAをクローニングし、その塩基配列を決定すれば、ATGpr [A. A. Salamov, T. Nishikawa, M. B. Swindells, Bioinformatics, 14: 384-390 (1998); <http://www.hri.co.jp/atgpr/>] 等の手法を用いて、それが全長cDNAクローンであるかどうかを評価することができる。しかし、これら公知の手法の組み合わせでは、確かにある程度の割合で全長cDNAを得ることができるものの、その効率においては改善の余地を残していた。そのため、発現頻度の低いmRNAについては、その全長cDNAをクローニングすることは依然として困難なことを考えられている。

【0006】

【発明が解決しようとする課題】

本発明は、ヒト由来の新規なタンパク質と、それをコードするDNA、並びにそれらの用途の提供を課題としている。

【0007】

【課題を解決するための手段】

我々は、オリゴキャップ法で作成した全長率の高いヒトcDNAライブラリーから、ATGpr 等で全長cDNAクローンであると予測される、ヒト全長cDNAを効率よくクローニングする方法を開発した。次いで、この方法で取得した全長率の高いcDNAクローンの塩基配列を5'側と3'側の両側から決定した。こうして得られた塩基配列を利用し、PSORT [K. Nakai & M. Kanehisa, Genomics, 14: 897-911 (1992)] でシグナル配列を持つと予測されるクローンを特異的に選別し、分泌タンパク質、または膜タンパク質をコードするcDNAを有すると予測できないクローンを取得した。

本発明における全長cDNAクローンは、[1]オリゴキャップ法による全長率の高

いcDNAライブラリーの作成、および[2]5'末端側の配列からの全長性の評価システムとの組み合わせによって取得することができた、より全長である確率の高いクローンである。

更に、この方法で取得したクローンの全長cDNA配列を解析し、その塩基配列がコードするアミノ酸配列を推定した。そして推定アミノ酸配列に基づいて、BLAST [S. F. Altschul, W. Gish, W. Miller, E. W. Myers & D. J. Lipman, J. Mol. Biol., 215: 403-410 (1990); W. Gish & D. J. States, Nature Genet., 3: 266-272 (1993); <http://www.ncbi.nlm.nih.gov/BLAST/>]によりGenBank (<http://www.ncbi.nlm.nih.gov/Web/Genbank/index.html>)やSwissProt (http://www.ebi.ac.uk/ebi_docs/swissprot_db/swisshome.html)を利用して相同性解析を行い本発明を完成した。

【0008】

すなわち本発明は、以下のタンパク質、このタンパク質をコードするDNA、並びにそれらの用途に関する。

まず本発明は、〔1〕配列番号：2、配列番号：4、配列番号：6、および配列番号：8から選択される、いずれかの配列番号として記載されたアミノ酸配列からなるタンパク質に関する。また本発明は、〔2〕〔1〕に記載のいずれかのタンパク質をコードするDNAに関する。本発明によるDNAの塩基配列は、たとえば〔3〕配列番号：1、配列番号：3、配列番号：5、および配列番号：7として記載された塩基配列のコード領域からなる。表1に、本発明による全長cDNAを有する実施例で単離したcDNAクローンの名称と、そのcDNAの塩基配列を表す配列番号、ならびにcDNAの塩基配列から推定されるアミノ酸配列を記載した配列番号の対応をまとめた。

【0009】

【表1】

アミノ酸配列	塩基配列	クローン名
配列番号：2	配列番号：1	PSEC0006

配列番号 : 4 配列番号 : 3 PSEC0043

配列番号 : 6 配列番号 : 5 PSEC0058

配列番号 : 8 配列番号 : 7 PSEC0211

【 0 0 1 0 】

更に本発明は、上記タンパク質やDNAに基づく以下の用途に関する。

〔 4 〕 〔 2 〕 または 〔 3 〕 に記載のDNAのいずれかが挿入されたベクター。

〔 5 〕 〔 2 〕 または 〔 3 〕 に記載のDNAのいずれかを発現可能に保持する形質転換体。

〔 6 〕 〔 5 〕 に記載の形質転換体を培養し、発現産物を回収する工程を含む、〔 1 〕 に記載のタンパク質のいずれかを製造する方法。

〔 7 〕 〔 2 〕 または 〔 3 〕 に記載のDNAのいずれか、またはその相補鎖にハイブリダイズするDNAであって、少なくとも 1 5 ヌクレオチドの鎖長を持つDNA。

〔 8 〕 〔 7 〕 に記載のDNAからなる、〔 1 〕 に記載されたタンパク質をコードするヒト全長cDNA合成用プライマー。

〔 9 〕 〔 7 〕 に記載のDNAからなる、〔 1 〕 に記載されたタンパク質をコードする遺伝子の検出用プローブ。

〔 1 0 〕 〔 2 〕 または 〔 3 〕 に記載のいずれかのDNAもしくはその一部に対するアンチセンスDNA。

〔 1 1 〕 次の工程を含む、〔 1 〕 に記載されたタンパク質をコードする全長 c D N A の合成方法。

a) c D N A ライブラリーを鋳型として 〔 8 〕 に記載のプライマーを起点とする相補鎖合成反応を行い、

b) 合成産物を回収する

〔 1 2 〕 c D N A ライブラリーが、オリゴキャップ法によって合成されたものである〔 1 1 〕 に記載の合成方法。

〔 1 3 〕 相補鎖の合成を P C R 法によって行う〔 1 1 〕 に記載の合成方法。

〔 1 4 〕 〔 1 〕 に記載のいずれかのタンパク質に対する抗体。

【 0 0 1 1 】

本発明において、ポリヌクレオチドとはヌクレオチドが多数重合した分子を意味する。重合するヌクレオチドの数は特に制限されないが、比較的重合度の低い場合には特にオリゴヌクレオチドとも表現する。本発明のポリヌクレオチド、またはオリゴヌクレオチドは、天然のものであることもできるし、化学的に合成されたものであることもできる。あるいはまた、鋳型となるDNAをもとにPCRのような酵素的な反応によって合成されたものであっても良い。

本発明によって提供されるcDNAはいずれも全長cDNAである。本発明における全長cDNAとは、そのcDNAの翻訳開始点となるATGコドンと終止コドンを備えたcDNAを意味する。したがって、天然のmRNAがタンパク質コード領域の上流や下流に本来備えている非翻訳領域の有無は問わない。

【 0 0 1 2 】

【発明の実施の形態】

本発明は、表 1 に示すように配列番号： 2、配列番号： 4、配列番号： 6、および配列番号： 8 のアミノ酸配列のいずれかからなる、ヒト由来のタンパク質である。表 2 に、本発明のタンパク質と、それをコードする全長cDNAクローンの特徴をまとめた。これらのクローンのうち、PSEC0058は、5'-末端配列がGenBankのdbESTの配列よりも長い。

【 0 0 1 3 】

【表 2】

クローン名	塩基数	アミノ酸数	翻訳開始	ATGpr1
			ATGの番号	
PSEC0006	1246bp	296aa	2	0.85
PSEC0043	1811bp	269aa	10	0.90
PSEC0058	4248bp	745aa	4	0.17
PSEC0211	1545bp	222aa	4	0.60

【 0 0 1 4 】

本発明のタンパク質は、そのアミノ酸配列が明らかとなっていることから、適当な発現系を適用して組み換え体として発現させることにより、あるいは、そのタンパクを特異的に認識する抗体を用いることで、その生物学的活性を解析することが可能である。

【0015】

例えば、次のようなことが考えられる。本発明のタンパク質を発現し、そのタンパク質を細胞内（各種培養細胞や初代培養細胞）にインジェクションし、細胞の変化を Ca^{2+} 等のシグナルの変動、細胞の生育状態の変化、あるいは、機能既知のタンパク質やmRNA等の変動等の解析によりクローン化した遺伝子の機能を解析することができる。また、本発明のタンパク質を特異的に認識する抗体を細胞内（各種培養細胞や初代培養細胞）にインジェクションし、細胞の変化を Ca^{2+} 等のシグナルの変動、細胞の生育状態の変化、あるいは、機能既知のタンパク質やmRNA等の変動等の解析によりクローン化した遺伝子の機能を解析することも考えられる。さらに、本発明のタンパク質を特異的に認識する抗体を用いることにより、細胞内での局在や詳細な組織での局在を解析することによる機能予測が考えられる。たとえば、胎児のホールボディ（ヒト胎児の入手が難しい場合はマウス等のアミノ酸レベルでの相同性が対応する遺伝子で一般的に高いのでマウス等の胎児でも解析可能である。特にサルは相同性が高い。）、各分化レベルでの細胞、培養細胞等での組織化学的解析により、クローン化した遺伝子の機能予測をすることが可能である。

【0016】

本発明のcDNAがコードしているタンパク質は、いずれも全長アミノ酸配列を備えることから、適当な発現系を適用して組み換え体として発現させることにより、あるいは、そのタンパクを特異的に認識する抗体を用いることで、前述したように、その生物学的活性を解析することが可能である。疾患と関連があるタンパク質であった場合には、タンパク質を発現して得られた特異認識抗体を用いて、特定の疾患とタンパク質の発現量や活性との相関を知ることができる。あるいは、ヒトの遺伝子と疾患のデータベースであるOnline Mendelian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/Omim/>)を利用した解析が可能である

。疾患関連タンパク質は、診断マーカー、発現・活性の増減を制御する薬剤、あるいは遺伝子治療のターゲットになるなど医薬品の開発等に有用である。そのうちでも、そのタンパク質が転写関連タンパク質やシグナル伝達関連タンパク質の場合には、疾患との関連が、藤井・田村・諸橋・影山・佐竹編の実験医学増刊「転写因子研究1999」Vol.17, No.3, (1999)や、遺伝子医学Vol.3, No.2(1999)で報告されていることより、医療産業上有用である。

【0017】

上記タンパク質を用いた機能の解析に基づいて、例えば以下のようにして医薬品開発を行うことができる。細胞の増殖・分化などの細胞状態を制御する因子である場合には、ある種の細胞に、本発明によって提供されるタンパク質や抗体を細胞内にマイクロインジェクションすることによって、細胞の増殖・分化などの細胞状態変化や、細胞内の特定の遺伝子の活性化または抑制を指標に低分子化合物等をスクリーニングすることができる。

このスクリーニングは、例えば、以下のように行うことができる。まず、本発明のタンパク質を発現させ組換えタンパク質の精製品を取得する。次いで、その精製タンパク質を、各種細胞株または初代培養細胞の細胞内にマイクロインジェクションして、増殖・分化などの細胞の変化を調べる。または、ある特定の細胞状態変化に作用することが知られている遺伝子の誘導をmRNA量、タンパク質量で検出する。あるいは、ある特定の細胞状態変化に作用することが知られている遺伝子産物（タンパク質）の働きにより変化した細胞内の物質（低分子化合物など）量で検出する。そのときに培養液等に活性をスクリーニングしたい物質（低分子でも高分子でも可能）を添加しておくことにより、細胞状態の変化が変わることを指標にスクリーニングできる。マイクロインジェクションしなくとも本発明で取得した遺伝子産物の変化で、スクリーニングできることもある。このようなスクリーニングにより、本発明によるタンパク質が細胞状態、機能を制御するのを活性化または抑制する物質が開発されれば、医薬品への応用が考えられる。

【0018】

さらに、具体的には次のような方法も可能である。まず、本発明のタンパク質を発現した形質転換細胞株を取得する。次いで、その形質転換細胞株と、もとの

未形質転換細胞株とにおいて、ある特定の遺伝子の変化をmRNA量、タンパク質量を検出する。あるいは、ある特定の遺伝子産物（タンパク質）の働きにより変化した細胞内の物質（低分子化合物など）量を検出する。さらには、ある特定の遺伝子の発現調節領域とマーカー遺伝子（ルシフェラーゼ、 β -ガラクトシダーゼ等）の融合遺伝子を導入した細胞に、本発明によって提供されるタンパク質を同時に発現させることによって、特定の遺伝子の発現の変化を、マーカー遺伝子産物（タンパク質）由来の活性で判定する。このようなスクリーニングにより、影響を受けたタンパク質や遺伝子が疾患に関連していた場合、本発明によるタンパク質を利用し、直接的に、または、間接的に、その発現や活性調節を行う化合物、遺伝子のスクリーニングが可能となる。

まず、本発明のタンパク質を発現させ組換えタンパク質の精製品を取得する。次に影響を受けたタンパク質や遺伝子を精製し、両者の結合を調べる。または、予め阻害剤の候補となる化合物を加えておいた後、それら結合の変化を調べる。あるいは本発明のタンパク質をコードする遺伝子の5'上流転写調節領域を取得し、それをマーカー遺伝子と融合した遺伝子を導入した細胞に、化合物などを添加して、当該遺伝子の発現を制御する因子を見いだす。このようなスクリーニングによって得られた化合物は、本発明によるタンパク質が関連した疾患に対して医薬品への応用が考えられる。スクリーニングによって得られた制御因子がタンパク質であっても、そのタンパク質の発現・活性に影響を与える化合物があれば、その化合物には本発明によるタンパク質が関連した疾患に対する医薬品への応用が考えられる。

【0019】

本発明によるタンパク質が酵素としての活性を有するとなれば、本発明によって提供されるタンパク質に適当な条件下で化合物を添加し、この化合物の変化を指標にスクリーニングすれば酵素活性の推測が可能である。また、この活性を指標に本発明によるタンパク質の活性を阻害する化合物のスクリーニングも可能である。このスクリーニングは、例えば、以下のように行うことが可能である。まず、本発明のタンパク質を発現させ組換えタンパク質の精製品を取得する。次いで、その精製タンパク質に、化合物を添加して、化合物量および反応生成物量を

調べる。または、予め阻害剤の候補となる化合物を加えておいた後、精製タンパク質と反応する化合物（基質）を加えて、その基質量および反応生成物量の変化を調べる。このようなスクリーニングにより、得られた化合物は、本発明のタンパク質が関連した疾患に対して、医薬品への応用が考えられる。

【0020】

本発明によるタンパク質を発現して得られた特異認識抗体を用いて、特定の疾患とタンパク質の発現量や活性との相関を知ることができる。あるいは、「Method in Molecular Biology」(Humana Press社)シリーズの『Molecular Diagnosis of Genetic Diseases』(Rob Elles編、1996)を参考に解析が可能である。

疾患関連タンパク質は、前述のようなスクリーニングの対象となり、その発現・活性を制御する薬剤の開発に有用である。また、関連した疾患の診断マーカー、あるいは遺伝子治療のターゲットになるなど医療産業上、有用である。

【0021】

以上により単離された化合物を医薬品として用いる場合には、単離された化合物自体を直接患者に投与する以外に、公知の製剤学的方法により製剤化して投与を行うことも可能である。例えば、薬理学上許容される担体もしくは媒体、具体的には、滅菌水や生理食塩水、植物油、乳化剤、懸濁剤などと適宜組み合わせて製剤化して投与することが考えられる。患者への投与は、例えば、動脈内注射、静脈内注射、皮下注射など当業者に公知の方法により行いうる。投与量は、患者の体重や年齢、投与方法などにより変動するが、当業者であれば適当な投与量を適宜選択することが可能である。また、該化合物がDNAによりコードされうるものであれば、該DNAを遺伝子治療用ベクターに組み込み、遺伝子治療を行うことも考えられる。投与量、投与方法は、患者の体重や年齢、症状などにより変動するが、当業者であれば適宜選択することが可能である。

【0022】

本発明のタンパク質は、組み換えタンパク質として、また天然のタンパク質として調製することが可能である。組み換えタンパク質は、例えば、後述するように本発明のタンパク質をコードするDNAを挿入したベクターを適当な宿主細胞に導入し、形質転換体内で発現したタンパク質を精製することにより調製すること

が可能である。一方、天然のタンパク質は、例えば、後述する本発明のタンパク質に対する抗体を結合したアフィニティーカラムを利用して調製することができる (Current Protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. Jhon Wiley & Sons Section 16.1-16.19)。アフィニティー精製に用いる抗体は、ポリクローナル抗体であってもモノクローナル抗体であってもよい。また、インビトロトランスレーション (例えば、「On the fidelity of mRNA translation in the nuclease-treated rabbit reticulocyte lysate system. Dasso, M.C., Jackson, R.J. (1989) NAR 17:3129-3144」参照) などにより本発明のタンパク質を調製することも可能である。

【0023】

本発明は、また、本発明のタンパク質の部分ペプチドを含む。また、抗体調製のための抗原ペプチドが含まれる。部分ペプチドが本発明のタンパク質に特異的であるためには、少なくとも7アミノ酸、好ましくは8アミノ酸以上、より好ましくは9アミノ酸以上のアミノ酸配列からなる。該部分ペプチドは、本発明のタンパク質に対する抗体や本発明のタンパク質の競合阻害剤の調製以外に、例えば、本発明のタンパク質に結合する受容体のスクリーニングなどに利用し得る。本発明の部分ペプチドは、例えば、遺伝子工学的手法、公知のペプチド合成法、あるいは本発明のタンパク質を適当なペプチダーゼで切断することによって製造する。

【0024】

また、本発明は、上記本発明のタンパク質をコードするDNAに関する。本発明のDNAとしては、本発明のタンパク質をコードしうるものであれば、その形態に特に制限はなく、cDNAの他、ゲノムDNA、化学合成DNAなども含まれる。また、本発明のタンパク質をコードしうる限り、遺伝暗号の縮重に基づく任意の塩基配列を有するDNAが含まれる。本発明のDNAは、上記のように、配列番号：1、配列番号：3、配列番号：5、および配列番号：7に示す塩基配列もしくはその一部をプローブとしたハイブリダイゼーション法やこれらDNA配列をもとに合成したプライマーを用いたPCR法等の常法により単離することが可能である。

たとえば実施例において単離した本発明による4クローンは新規で、かつ全長

cDNAである。本発明に基づくすべてのcDNAクローンは、次のように特徴付けることもできる。

【0025】

すなわち、オリゴキャップ法で取得された全長性の高いcDNAであり、その5'末端配列データの特徴をもとに、5'末端の全長性を予測するATGpr（あるいはATGpr 1と記載している）のスコアにより選別されており、さらにシグナル配列の存在を予測するPSORTにより5'末端にシグナル配列を見出すことができず、更にタンパクコーディング領域中に膜貫通領域を有さないものが選別されている。また、選別されたクローンは5'末端配列の相同性検索によりヒトmRNAに対して同一でない（すなわち新規である）ことがわかっている。

【0026】

また、本発明は、本発明のDNAが挿入されたベクターに関する。本発明のベクターとしては、挿入したDNAを安定に保持するものであれば特に制限されず、例えば宿主に大腸菌を用いるのであれば、クローニング用ベクターとしてはpBlue scriptベクター（Stratagene社製）などが好ましい。本発明のタンパク質を生産する目的においてベクターを用いる場合には、特に発現ベクターが有用である。発現ベクターとしては、試験管内、大腸菌内、培養細胞内、生物個体内でタンパク質を発現するベクターであれば特に制限されないが、例えば、試験管内発現であればpBESTベクター（プロメガ社製）、大腸菌であればpETベクター（Invitrogen社製）、培養細胞であればpME18S-FL3ベクター（GenBank Accession No. AB009864）、生物個体であればpME18Sベクター（Mol Cell Biol. 8:466~472(1988)）などが好ましい。ベクターへの本発明のDNAの挿入は常法により制限酵素サイトを用いたりガーゼ反応により行うことができる（Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 11.4~11.11）。

【0027】

また、本発明は、本発明のベクターを保持する形質転換体に関する。本発明のベクターが導入される宿主細胞としては特に制限はなく、目的に応じて種々の宿主細胞が用いられる。タンパク質を高発現させるための真核細胞としては、例え

ば、COS細胞、CHO細胞などを例示することができる。

【0028】

宿主細胞へのベクター導入は、例えば、リン酸カルシウム沈殿法、電気パルス穿孔法 (Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 9.1-9.9)、リポフェクタミン法 (G IBCO-BRL社製)、マイクロインジェクション法などの方法で行うことが可能である。

【0029】

また本発明は、本発明のタンパク質をコードする配列番号：1、配列番号：3、配列番号：5、および配列番号：7に記載の塩基配列からなるDNA、またはその相補鎖と特異的にハイブリダイズし、少なくとも15ヌクレオチドの鎖長を有するDNAに関する。ここで「特異的にハイブリダイズする」とは、通常のハイブリダイゼーション条件下、好ましくは厳格な条件下で、本発明のタンパク質をコードする配列番号：1、配列番号：3、配列番号：5、および配列番号：7に記載のDNA、またはその相補鎖とハイブリダイズし、他のタンパク質をコードするDNAとはハイブリダイズしないことを意味する。このようなDNAは、本発明のDNAを検出、単離するためのプローブとして、また、本発明のDNAを増幅するためのプライマーとして利用することが可能である。プライマーとして用いる場合には、通常、15bp～100bp、好ましくは15bp～35bpの鎖長を有する。また、プローブとして用いる場合には、本発明のDNAの少なくとも一部若しくは全部の配列を有し、少なくとも15bpの鎖長のDNAが用いられる。

【0030】

本発明のDNAは、本発明のタンパク質の異常を検査・診断するために利用できる。例えば、本発明のDNAをプローブやプライマーとして用いたノーザンハイブリダイゼーションやRT-PCRにより、発現異常を検査したり、本発明のDNAをプライマーとして用いたポリメラーゼ連鎖反応(PCR)によりゲノムDNA-PCRやRT-PCRにより本発明のタンパク質をコードするDNAやその発現制御領域を増幅し、RFLP解析、SSCP、シーケンシング等の方法により、配列の異常を検査・診断することができる。

【0031】

また、「配列番号：1、配列番号：3、配列番号：5、および配列番号：7に記載の塩基配列からなるDNA、またはその相補鎖と特異的にハイブリダイズし、少なくとも15ヌクレオチドの鎖長を有するDNA」には、本発明のタンパク質の発現を抑制するためのアンチセンスDNAが含まれる。アンチセンスDNAは、アンチセンス効果を引き起こすために、少なくとも15bp以上、好ましくは100bp、さらに好ましくは500bp以上の鎖長を有し、通常、3000bp以内、好ましくは2000bp以内の鎖長を有する。このようなアンチセンスDNAには、本発明のタンパク質の異常（機能異常や発現異常）などに起因した疾患の遺伝子治療への応用も考えられる。該アンチセンスDNAは、例えば、本発明のタンパク質をコードするDNA（例えば、配列番号：1、配列番号：3、配列番号：5、および配列番号：7に記載のDNA）の配列情報を基にホスホロチオネート法（Stein, 1988 Physicochemical properties of phosphorothioate oligodeoxynucleotides, Nucleic Acids Res 16, 3209-21 (1988)) などにより調製することが可能である。

【0032】

本発明のDNAまたはアンチセンスDNAは、遺伝子治療に用いる場合には、例えば、レトロウイルスベクター、アデノウイルスベクター、アデノ随伴ウイルスベクターなどのウイルスベクターやリポソームなどの非ウイルスベクターなどを利用して、ex vivo法やin vivo法などにより患者へ投与を行う。

【0033】

また、本発明は、本発明のタンパク質に結合する抗体に関する。本発明の抗体の形態には特に制限はなく、ポリクローナル抗体やモノクローナル抗体または抗原結合性を有するそれらの一部も含まれる。また、全てのクラスの抗体が含まれる。さらに、本発明の抗体には、ヒト化抗体などの特殊抗体も含まれる。

【0034】

本発明の抗体は、ポリクローナル抗体の場合には、常法に従いアミノ酸配列に相当するオリゴペプチドを合成して家兎に免疫することにより得ることが可能であり（Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 11.12~11.13）、一方、モノクローナ

ル抗体の場合には、常法に従い大腸菌で発現し精製したタンパク質を用いてマウスを免疫し、脾臓細胞と骨髓腫細胞を細胞融合させたハイブリドーマ細胞の中から得ることができる (Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 11.4~11.11)。

【0035】

本発明のタンパク質に結合する抗体は、本発明のタンパク質の精製に加え、例えば、本発明のタンパク質の発現異常や構造異常の検査・診断に利用することも考えられる。具体的には、例えば組織、血液、または細胞などからタンパク質を抽出し、ウェスタンブロッティング、免疫沈降、ELISA等の方法による本発明のタンパク質の検出を通して、発現や構造の異常の有無を検査・診断することができる。

【0036】

また、本発明のタンパク質に結合する抗体を、本発明のタンパク質に関連した疾患の治療などの目的に利用することも考えられる。抗体を患者の治療目的で用いる場合には、ヒト抗体またはヒト化抗体が免疫原性の少ない点で好ましい。ヒト抗体は、免疫系をヒトのものと入れ換えたマウス (例えば、「Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice, Mendez, M.J. et al. (1997) Nat. Genet. 15:146-156」参照) に免疫することにより調製することができる。また、ヒト化抗体は、モノクローナル抗体の超可変領域を用いた遺伝子組み換えによって調製することができる (Methods in Enzymology 203, 99-121 (1991))。

【0037】

【実施例】

次に、本発明を実施例によりさらに具体的に説明するが、本発明は下記実施例に限定されるものではない。

実施例 1. オリゴキャップ法による cDNA ライブラリーの作製

ヒト胎児精巣由来のテラトカルシノーマ細胞でレチノイン酸処理により神経細胞に分化可能な NT-2 神経前駆細胞 (Stratagene 社より購入) を用いた。添付マニュアルに従って、次の条件で培養細胞を調製した。

- (1) NT-2細胞をレチノイン酸で誘導しないで培養 (NT2RM1)
- (2) NT-2細胞を培養後、レチノイン酸を添加して誘導後、48時間培養 (NT2RP1)
- (3) NT-2細胞を培養後、レチノイン酸を添加して誘導後、2週間培養 (NT2RP2)

これらの培養細胞をそれぞれ集めて、文献 (J. Sambrook, E. F. Fritsch & T. Maniatis, Molecular Cloning Second edition, Cold Spring harbor Laboratory Press 1989) 記載の方法により mRNA を抽出した。さらに、オリゴdTセルロースで poly(A)⁺RNA を精製した。

同様に、ヒト胎児より脳を多く含む組織 (HEMBA1) より、文献 (J. Sambrook, E. F. Fritsch & T. Maniatis, Molecular Cloning Second edition, Cold Spring harbor Laboratory Press, 1989) 記載の方法により mRNA を抽出した。さらに、オリゴdTセルロースで poly(A)⁺RNA を精製した。

【0038】

それぞれの poly(A)⁺RNA よりオリゴキャップ法 [M. Maruyama and S. Sugano, Gene, 138: 171-174 (1994)] により cDNA ライブラリーを作成した。Oligo-cap linker (配列番号: 9) およびオリゴdTプライマー (配列番号: 10) を用いて文献 [鈴木・菅野, 蛋白質 核酸 酵素, 41: 197-201 (1996)、Y. Suzuki et al., Gene, 200: 149-156 (1997)] の記載にしたがって BAP (Bacterial Alkaline Phosphatase) 処理、TAP (Tobacco Acid Phosphatase) 処理、RNA ライゲーション、第一鎖 cDNA の合成と RNA の除去を行った。次いで、5' (配列番号: 11) と 3' (配列番号: 12) の PCR プライマーを用い PCR (polymerase chain reaction) により 2 本鎖 cDNA に変換し、SfiI 切断した。次いで、DraIII で切断したベクター-pUC19FL3 (図 1) (NT2RM1 と NT2RP1) または pME18SFL3 (図 1) (GenBank AB009864, Expression vector) (NT2RP2、HEMBA1) に cDNA の方向性を決めてクローニングし、cDNA ライブラリーを作成した。これらより得たクローンのプラスミド DNA について、挿入 cDNA サイズが 1 kb 以下のクローンを除いた後、cDNA の 5' 端と 3' 端の塩基配列を DNA シーケンシング試薬 (Dye Terminator Cycle Sequencing FS Ready Reaction Kit, dRhodamine Terminator Cycle Sequencing FS Ready Reaction

KitまたはBigDye Terminator Cycle Sequencing FS Ready Reaction Kit, PE Biosystems社製)を用い、マニュアルに従ってシーケンシング反応後、DNAシーケンサー (ABI PRISM 377, PE Biosystems社製) でDNA塩基配列を解析した。

【0039】

NT2RP2とHEMBA1のオリゴキャップ高全長率cDNAライブラリーは、真核細胞での発現が可能な発現ベクターpME18SFL3を用いて作製した。pME18SFL3にはクロニング部位の上流にSR α プロモーターとSV40 small tイントロンが組み込まれており、またその下流にはSV40ポリA付加シグナル配列部位が挿入されている。pME18SFL3のクローン化部位は非対称性のDraIIIサイトとなっており、cDNA断片の末端にはこれと相補的なSfiI部位を付加しているので、クローン化したcDNA断片はSR α プロモーターの下流に一方向性に挿入される。したがって、全長cDNAを含むクローンでは、得られたプラスミドをそのままCOS細胞に導入することにより、一過的に遺伝子を発現させることが可能である。すなわち、非常に容易に、遺伝子産物である蛋白質として、あるいはそれらの生物学的活性として実験的に解析することが可能となっている。

【0040】

オリゴキャップ法で作製したライブラリーのcDNAの5'-末端の全長率を次の方法で、求めた。公共データベース中のヒト既知mRNAと5'-末端配列が一致する全クローンについて、公共データベース中の既知mRNA配列より長く5'-末端が伸びている場合と5'-末端は短いが翻訳開始コドンは有している場合を「全長」と判断し、翻訳開始コドンを含んでいない場合を「非全長」と判断した。各ライブラリーでのcDNAクローンの5'-末端の全長率 [全長クローン数 / (全長クローン数 + 非全長クローン数)] をヒト既知mRNAと比較することにより求めた (NT2RM1 : 69% ; NT2RP1 : 75% ; NT2RP2 : 62% ; HEMBA1 : 53%)。この結果より、5'-端配列の全長率が非常に高いことが分かった。

cDNAライブラリーとクローンとの関係は次のとおりである。

NT2RM1 : PSEC0006

NT2RP1 : PSEC0043

NT2RP2 : PSEC0058

【0041】

実施例 2. ATGprとESTiMateFLでのcDNAの5'-末端の全長率の評価

ATGpr は、ATGコドンの周辺の配列の特徴から翻訳開始コドンであるかどうかを予測するためにヘリックス研究所のA. A. Salamov, T. Nishikawa, M. B. Swindellsにより開発したプログラムである。結果は、そのATGが真の開始コドンである期待値（以下、ATGpr1と記載することもある）で表した（0.05-0.94）。尚、このプログラムのcDNAの5'-末端であるかどうかを考慮しない場合の解析結果の感度と特異性はともに66%と評価している。一方、このプログラムを全長率65%のオリゴキャップ法で作製したライブラリーからのcDNAクローンの5'-末端配列に適用してATGpr1値を0.6以上でクローンを選択した場合、全長クローン（ORFのN-末端まで持つクローン）評価の感度と特異性はともに82~83%まで上昇した。さらに、このプログラムをオリゴキャップ法で作製したヒトcDNAライブラリーからのクローンの5'-末端配列が既知ヒトmRNAと一致する17,365クローンについて評価した結果を示す。方法は、既知ヒトmRNAと一致するクローンについてATGpr1の最大値をだす。次いで、各クローンの5'-末端配列を既知ヒトmRNAのORFと比較し、全長か非全長かを決めた。その結果をまとめたものを表3に示した。ATGprとオリゴキャップ法で作製したヒトcDNAライブラリーからのクローンの組み合わせによる選択が、非常に有効なことを示している。

【0042】

【表3】

ATGpr1の 最大値	(全長+非全長) の数	全長の数	全長率
>=0.70	10,226	8,428	82.4%
>=0.50	12,171	9,422	77.4%
>=0.30	14,102	10,054	71.3%
>=0.17	15,647	10,385	66.4%

≥ 0.05 17,365 10,608 61.1%

* * 全長の数：ORFのN-末端までもつクローンの数；非全長の数：ORFのN-末端まで持っていないクローンの数；全長率：全長の数／（全長＋非全長）の数

【0043】

ESTiMateFLは、公共データベース中のESTの5'-末端配列や3'-末端配列との比較による全長cDNAの可能性の高いクローンを選択するヘリックス研究所の西川・太田らにより開発された方法である。

この方法は、あるcDNAクローンの5'-末端や3'-末端配列よりも、長く伸びたESTが存在する場合には、そのクローンは「全長ではない可能性が高い」と判断する方法で、大量処理可能なようにシステム化したものである。公共データベース中のEST配列より長く5'-末端が伸びている場合、および5'-末端が短いクローンでも両者の差が50塩基以内である場合を便宜的に全長とし、それ以上短い場合を非全長とした。ESTとの比較による完全長らしさの評価では、比較対照とするESTの数が多ければ予測精度は高まるが、対象ESTが少ない場合には予測結果の信頼性が低くなる欠点はある。この方法は、5'-末端配列での全長率が約60%のオリゴキャップ法によるcDNAクローンから全長ではない可能性の高いクローンを排除するのに使えば有効である。また、ESTiMateFLは、公共データベースへのEST登録が適当数あるヒト未知mRNAのcDNAの3'-末端配列の全長性を評価するには、特に有効な方法である。

【0044】

その結果をまとめたものを表4、および表5に示した。ATGprとESTiMateFLの両プログラムを組み合わせて、オリゴキャップ法で作製したヒトcDNAライブラリーからのクローンの5'-末端配列の全長性を評価すれば、ATGprの値の低いクローンでも全長率が高くなることが確認された。この結果を全長cDNA配列を決めたクローンの5'-末端配列の全長性の評価に応用した。なお表中、全長の数とはORFのN-末端までもつクローンの数、非全長の数とはORFのN-末端まで持っていないクローンの数、そして全長率とは全長の数／（全長＋非全長）の数を意味する。

【0045】

【表 4】

既知ヒト mRNA の ORF と比較し、全長であると判定したオリゴキャップ法で取得した cDNA クローンの 5'-末端配列の EST に対する全長率

ATGpr1 の 最大値	(全長+非全長) の数	非全長 の数	全長率
>=0.30	8,646	907	90.5%
>=0.17	10,158	1,150	89.8%
>=0.05	15,351	2,728	84.9%

【0046】

【表 5】

既知ヒト mRNA の ORF と比較し、非全長であると判定したオリゴキャップ法で取得した cDNA クローンの 5'-末端配列の EST に対する全長率

ATGpr1 の 最大値	(全長+非全長) の数	非全長 の数	全長率
>=0.30	1,271	2,156	37.1%
>=0.17	1,678	2,907	36.6%
>=0.05	2,512	4,529	35.7%

【0047】

実施例 3. 全長率の高いクローンの選択と全長配列解析

本発明によるクローンのうち、PSEC0006-PSEC0058 については、5'-端配列データでの ORF (アミノ酸翻訳領域) が存在するものを選別したクローンであり、5'-端配列データ (one pass sequencing) の ATGpr に基づく選別はしていない。また、PSEC0211 については、5'-端配列データ (one pass sequencing) から ATGpr1 の

最大値が0.7以上で、5'-端配列データでのORF（アミノ酸翻訳領域）が存在するものを選別したクローンである。

選択した4クローンについて、全長cDNAの塩基配列、並びに推定アミノ酸配列を決定した。塩基配列は、次に示す3種の方法を組み合わせ、各方法によって決定した塩基配列を完全にオーバーラップさせ、最終的な確定塩基配列を決定した。決定されたcDNA配列から、推定アミノ酸配列を明らかにした。それらの結果を配列表に示した。

(1) Licor DNAシーケンサーを用いたcDNA挿入断片両末端からのロングリードシーケンス（Licorシーケンサー（アロカ社販売）のマニュアルに従ってシーケンシング反応後、LicorシーケンサーでDNA塩基配列を解析した）、

(2) AT2トランスポゾン試験管内転移を用いたPrimer Island法によるネステッドシーケンス[S. E. Devine and J. D. Boeke, Nucleic Acids Res., 22: 3765-3772, (1994)]（PE Biosystems社製のキットとマニュアルにしたがってクローンを取得後、PE Biosystems社製のDNAシーケンシング試薬でマニュアルに従ってシーケンシング反応し、ABI PRISM 377でDNA塩基配列を解析した）

(3) カスタム合成DNAプライマーを用いたダイデオキシターミネーター法によるプライマーウォーキング（カスタム合成DNAプライマーをもちいPE Biosystems社製のDNAシーケンシング試薬でマニュアルに従ってシーケンシング反応し、ABI PRISM 377でDNA塩基配列を解析した）

それらの配列について、ATGprによる解析およびGenBankやSwissProtに対するBLAST解析を行った。本発明の4クローンと高度な相同性を示す公知のアミノ酸配列は確認できなかった。また、特徴的なモチーフを見出すこともできなかった。

【0048】

以上の結果から、これら4クローン（PSEC0006, PSEC0043, PSEC0058, PSEC0211）は、「全長率の高いオリゴキャップ法で作成したヒトcDNAライブラリーから、ATGpr等で全長cDNAクローンであると予測されるクローン」と定義される。ただし本発明によるクローンの中でATGpr1の値が低いPSEC0058（ATGpr1 0.17）は、cDNA5'-末端配列データ（one pass sequencing）の5'-末端配列データに基づくORFが長いものとして選別してきたクローンについて全長配列解析したもので

ある。つまりPSEC0058は、ATGprに基づいて選別したものではない。PSEC0058についてESTとの比較をした結果、ESTよりも長いクローンであると判定された。

【0049】

【発明の効果】

本発明により、4種の新規なタンパク質と、それをコードする全長cDNAが提供された。全長cDNAの分離が進んでいないヒトにおいて、新規な全長cDNAを提供した意義は大きい。本発明の全長cDNAはヒト由来であるので、疾患に関連している可能性がある。疾患に関連した遺伝子やタンパク質は、診断マーカーとして有用である。また発現や活性を制御する化合物の探索、あるいは遺伝子治療の標的となるなど、医薬品開発などに有用である。

【0050】

【配列表】

SEQUENCE LISTING

<110> Helix Research Institute

<120> Full length cDNAs and proteins.

<130> H1-108

<140>

<141>

<160> 12

<170> PatentIn Ver. 2.0

<210> 1

<211> 1246

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (264)..(1151)

<400> 1

agacgggcgg cgcgtggcgg aaggcaggct tgctcctcgg ggtgggggag ggtatccggc 60

ttaagggggc tgcggtggac accacttctt aatgtcgggg gtcttcgcgg cgctcacctc 120

ggctcctagg gttcgggacg gtacgcacca gccaccttcg cgccgaaggc ggtagggcgc 180

cacggagagg aaccgctcta ggcacgtaag gcctcgtgag gtigcgtcgc gcgcggagca 240

ctctgggact tgtagttctg gag atg gag cga gct gtg ccg ctc gcg gtg cct 293

Met Glu Arg Ala Val Pro Leu Ala Val Pro

1

5

10

ctg ggt cag aca gag gtg ttc cag gcc ttg cag cgg ctc cat atg acc 341

Leu Gly Gln Thr Glu Val Phe Gln Ala Leu Gln Arg Leu His Met Thr

15

20

25

atc ttc tcc cag agc gtc tca cca tgt ggg aag ttt ctg gcg gct ggc 389

Ile Phe Ser Gln Ser Val Ser Pro Cys Gly Lys Phe Leu Ala Ala Gly

30

35

40

aac aat tac ggg cag att gcc atc ttc agc ttg tcc tct gct ttg agc 437

Asn Asn Tyr Gly Gln Ile Ala Ile Phe Ser Leu Ser Ser Ala Leu Ser

45

50

55

tca gaa gcc aaa gag gaa agt aag aag ccg gtg gtg act ttc caa gcc 485

Ser Glu Ala Lys Glu Glu Ser Lys Lys Pro Val Val Thr Phe Gln Ala

60

65

70

cat gat ggg ccc gtc tat agc atg gtt tcc acc gat cga cat ctg ctt 533

His Asp Gly Pro Val Tyr Ser Met Val Ser Thr Asp Arg His Leu Leu

75

80

85

90

agt gct ggg gat ggg gag gag aag gcc tgg ctt tgg gcg gag atg ctc 581

Ser Ala Gly Asp Gly Glu Glu Lys Ala Trp Leu Trp Ala Glu Met Leu

95

100

105

aag aag ggc tgt aag gag ctg tgg cgt cgt cag cct cca tac agg acc 629

Lys Lys Gly Cys Lys Glu Leu Trp Arg Arg Gln Pro Pro Tyr Arg Thr

110

115

120

agc ctg gaa gtg cct gag atc aac gct ttg ctg ctg gtc ccc aag gag 677

Ser Leu Glu Val Pro Glu Ile Asn Ala Leu Leu Leu Val Pro Lys Glu

125

130

135

aat tcc ctc atc ctg gct ggg gga gac tgt cag ttg cac act atg gac 725

Asn Ser Leu Ile Leu Ala Gly Gly Asp Cys Gln Leu His Thr Met Asp

140

145

150

ctt gaa act ggg act ttc acg agg gtc ctc cgg ggc cac aca gac tac 773

Leu Glu Thr Gly Thr Phe Thr Arg Val Leu Arg Gly His Thr Asp Tyr

155

160

165

170

atc cac tgc ctg gca ctg cgg gaa agg agc cca gag gtg ctg tca ggt 821

Ile His Cys Leu Ala Leu Arg Glu Arg Ser Pro Glu Val Leu Ser Gly

175

180

185

ggc gag gat gga gct gtt cga ctt tgg gac ctg cgc aca gcc aag gag 869

Gly Glu Asp Gly Ala Val Arg Leu Trp Asp Leu Arg Thr Ala Lys Glu

190

195

200

gtc cag acg atc gag gtc tat aag cac gag gag tgc tcg agg ccc cac 917

Val Gln Thr Ile Glu Val Tyr Lys His Glu Glu Cys Ser Arg Pro His

205

210

215

aat ggg cgc tgg att gga tgt ttg gca act gat tcc gac tgg atg gtc 965

Asn Gly Arg Trp Ile Gly Cys Leu Ala Thr Asp Ser Asp Trp Met Val

220

225

230

tgt gga ggg ggc cca gcc ctc acc ctc tgg cac ctc cga tcc tcc aca 1013

Cys Gly Gly Gly Pro Ala Leu Thr Leu Trp His Leu Arg Ser Ser Thr

235

240

245

250

ccc acc acc atc ttc ccc atc cgg gcg cca cag aag cac gtc acc ttc 1061

Pro Thr Thr Ile Phe Pro Ile Arg Ala Pro Gln Lys His Val Thr Phe

255

260

265

tac cag gac ctg gtc ctg aca gct gca ggc aac agc tgc cgg gtg gat 1109

Tyr Gln Asp Leu Val Leu Thr Ala Ala Gly Asn Ser Cys Arg Val Asp

270

275

280

gtc ttc acc aac ctg ggt tac cga gcc ttc tcc ctg tcc ttc 1151

Val Phe Thr Asn Leu Gly Tyr Arg Ala Phe Ser Leu Ser Phe

285

290

295

tgatctctga cgacaccccc agccagctca gggttttaga gtgtttttca ttttcttttt 1211

tttttttttt ttacaataaa gtttcaggct ttttt 1246

<210> 2

<211> 296

<212> PRT

<213> Homo sapiens

<400> 2

Met Glu Arg Ala Val Pro Leu Ala Val Pro Leu Gly Gln Thr Glu Val

1

5

10

15

Phe Gln Ala Leu Gln Arg Leu His Met Thr Ile Phe Ser Gln Ser Val

20

25

30

Ser Pro Cys Gly Lys Phe Leu Ala Ala Gly Asn Asn Tyr Gly Gln Ile

35

40

45

Ala Ile Phe Ser Leu Ser Ser Ala Leu Ser Ser Glu Ala Lys Glu Glu

50

55

60

Ser Lys Lys Pro Val Val Thr Phe Gln Ala His Asp Gly Pro Val Tyr

65

70

75

80

Ser Met Val Ser Thr Asp Arg His Leu Leu Ser Ala Gly Asp Gly Glu

85

90

95

Glu Lys Ala Trp Leu Trp Ala Glu Met Leu Lys Lys Gly Cys Lys Glu

100

105

110

Leu Trp Arg Arg Gln Pro Pro Tyr Arg Thr Ser Leu Glu Val Pro Glu

115

120

125

Ile Asn Ala Leu Leu Leu Val Pro Lys Glu Asn Ser Leu Ile Leu Ala

130

135

140

Gly Gly Asp Cys Gln Leu His Thr Met Asp Leu Glu Thr Gly Thr Phe

145

150

155

160

Thr Arg Val Leu Arg Gly His Thr Asp Tyr Ile His Cys Leu Ala Leu

165

170

175

Arg Glu Arg Ser Pro Glu Val Leu Ser Gly Gly Glu Asp Gly Ala Val

180

185

190

Arg Leu Trp Asp Leu Arg Thr Ala Lys Glu Val Gln Thr Ile Glu Val

195

200

205

Tyr Lys His Glu Glu Cys Ser Arg Pro His Asn Gly Arg Trp Ile Gly

210

215

220

Cys Leu Ala Thr Asp Ser Asp Trp Met Val Cys Gly Gly Gly Pro Ala
225 230 235 240

Leu Thr Leu Trp His Leu Arg Ser Ser Thr Pro Thr Thr Ile Phe Pro
245 250 255

Ile Arg Ala Pro Gln Lys His Val Thr Phe Tyr Gln Asp Leu Val Leu
260 265 270

Thr Ala Ala Gly Asn Ser Cys Arg Val Asp Val Phe Thr Asn Leu Gly
275 280 285

Tyr Arg Ala Phe Ser Leu Ser Phe
290 295

<210> 3

<211> 1811

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (659)..(1465)

<400> 3

agtgcctgcg gccctcggcg gcctagtaca cacgcacctg agtgagtggc accagaggac 60

cctctccatg tttagggacc tcctgggcct caggagcgtg gcgcccggcc ctgggcggac 120

tccccccatc cgcgggcgcg aatggtccgg gtcgcgtccg cagtgtgtgt ggctgtctccc 180

tggttgctgg gtgcaaagtg ctgggttctg ggtttctgga ttcgcgggcc gttcacacgt 240

agcctgtgcc ggctcctcgg gtgagtcctg ccgcgcgcgg tgccccggga cggcctaggc 300

tgcggggggt ccggggcccc aggcattccg ggctgcagat tgacggggat cccggatgca 360

ccgcgcgccc ccgcgcccctc accgacgggt ccagacctgg tgggaagaag gtgcgggggac 420

gggtccctga ggatcccgat gcctacgagc caagatgctc agctttatag gtgtgacct 480

cacatgtgac ttcaccicag ttttgtgatc cgtaaaatgg acaaattcga agctacttca 540

cagtgtgtgt gagaggatta aatgaaacaa tgcttgtaaa gctctttgca ggaggagacc 600

tcggaagcag ggcctggccg gcagagcaca cctgtgtgtca ccagggacca caggcagc 658

atg aag acc ccc gtg gag ctg gcc gtc agt ggg atg cag acc ctc ggc 706

Met Lys Thr Pro Val Glu Leu Ala Val Ser Gly Met Gln Thr Leu Gly

1 5 10 15

ctt cag cac cgc tgc cga ggt ggc tac cgg gtc aag gcc agg acg tca 754

Leu Gln His Arg Cys Arg Gly Gly Tyr Arg Val Lys Ala Arg Thr Ser

20 25 30

tat gtg gat gag act ctg ttt ggc agc cca gca ggc acc cgg cct acc 802

Tyr Val Asp Glu Thr Leu Phe Gly Ser Pro Ala Gly Thr Arg Pro Thr

35

40

45

cca ccg gac ttc gat ccg ccc tgg gtg gag aag gct aac aga acc aga 850

Pro Pro Asp Phe Asp Pro Pro Trp Val Glu Lys Ala Asn Arg Thr Arg

50

55

60

ggc gtg ggc aag gag gca tcg aag gcc ttg ggg gca aag ggg agc tgt 898

Gly Val Gly Lys Glu Ala Ser Lys Ala Leu Gly Ala Lys Gly Ser Cys

65

70

75

80

gag acc acc ccc tca agg ggc agc acc ccc acc ctc aca cca agg aag 946

Glu Thr Thr Pro Ser Arg Gly Ser Thr Pro Thr Leu Thr Pro Arg Lys

85

90

95

aag aac aaa tac aga ccc atc agc cac acc ccg tct tac tgt gat gag 994

Lys Asn Lys Tyr Arg Pro Ile Ser His Thr Pro Ser Tyr Cys Asp Glu

100

105

110

tcg ctg ttt ggc tcc cga tct gaa ggc gcc agc ttc ggg gcc ccg cgg 1042

Ser Leu Phe Gly Ser Arg Ser Glu Gly Ala Ser Phe Gly Ala Pro Arg

115

120

125

atg gcg aag ggg gat gcc gca aag ctc cgt gct ctc ttg tgg acg cca 1090

Met Ala Lys Gly Asp Ala Ala Lys Leu Arg Ala Leu Leu Trp Thr Pro

130

135

140

cca cct acc ccc agg ggt agc cac tcg ccc cgc ccc agg gag gca cca 1138

Pro Pro Thr Pro Arg Gly Ser His Ser Pro Arg Pro Arg Glu Ala Pro

145

150

155

160

ctg cga gcc att cac cca gct ggt ccc tcc aag aca gag ccg ggg cca 1186

Leu Arg Ala Ile His Pro Ala Gly Pro Ser Lys Thr Glu Pro Gly Pro

165

170

175

gcg gca gac tcc cag aag tta tct atg ggt ggg tta cac tct tca cgc 1234

Ala Ala Asp Ser Gln Lys Leu Ser Met Gly Gly Leu His Ser Ser Arg

180

185

190

ccc ctg aag cgg gga ctt tcc cat tcc ctc acc cac ctg aat gtc ccc 1282

Pro Leu Lys Arg Gly Leu Ser His Ser Leu Thr His Leu Asn Val Pro

195

200

205

agc act ggt cat cca gcc acc agt gcc ccc cac aca aat ggg cct cag 1330

Ser Thr Gly His Pro Ala Thr Ser Ala Pro His Thr Asn Gly Pro Gln

210

215

220

gat ctc agg cct tcc acg tca ggg gtg acc ttc cgg agc ccc ctg gtg 1378

Asp Leu Arg Pro Ser Thr Ser Gly Val Thr Phe Arg Ser Pro Leu Val

225

230

235

240

act tcc agg gct cgc tca gtt agc att tca gtg cca tct acc cca cga 1426

Thr Ser Arg Ala Arg Ser Val Ser Ile Ser Val Pro Ser Thr Pro Arg

245

250

255

cga ggt ggg gcc acc cag aaa cca aag ccc cct tgg aaa tgataactctt 1475

Arg Gly Gly Ala Thr Gln Lys Pro Lys Pro Pro Trp Lys

260

265

tcattcagggt tgcctatggg gccacggcga caggtatggc cccttgccag ggtaggagga 1535

cattcatcac ccagggaacc ccaggtatta aagaagcccc tgtgggggca gacagacata 1595

gcaggggtgg gcagtgcctc cttttatcct gacaatctct agtcgattct tgcctttttc 1655

tcccgattgc ggatttgggg gccacctcta agatgcctct ctccagccct gtctcaacca 1715

tactccaaat tagtgccaac ccaggggcct ggcacctccc acatcatcca ttgtcttgct 1775

gccaaagtgcg aataaacggc gtgattgcc aacctgg 1811

<210> 4

<211> 269

<212> PRT

<213> Homo sapiens

<400> 4

Met Lys Thr Pro Val Glu Leu Ala Val Ser Gly Met Gln Thr Leu Gly

1 5 10 15

Leu Gln His Arg Cys Arg Gly Gly Tyr Arg Val Lys Ala Arg Thr Ser

20 25 30

Tyr Val Asp Glu Thr Leu Phe Gly Ser Pro Ala Gly Thr Arg Pro Thr

35 40 45

Pro Pro Asp Phe Asp Pro Pro Trp Val Glu Lys Ala Asn Arg Thr Arg

50

55

60

Gly Val Gly Lys Glu Ala Ser Lys Ala Leu Gly Ala Lys Gly Ser Cys

65

70

75

80

Glu Thr Thr Pro Ser Arg Gly Ser Thr Pro Thr Leu Thr Pro Arg Lys

85

90

95

Lys Asn Lys Tyr Arg Pro Ile Ser His Thr Pro Ser Tyr Cys Asp Glu

100

105

110

Ser Leu Phe Gly Ser Arg Ser Glu Gly Ala Ser Phe Gly Ala Pro Arg

115

120

125

Met Ala Lys Gly Asp Ala Ala Lys Leu Arg Ala Leu Leu Trp Thr Pro

130

135

140

Pro Pro Thr Pro Arg Gly Ser His Ser Pro Arg Pro Arg Glu Ala Pro

145

150

155

160

Leu Arg Ala Ile His Pro Ala Gly Pro Ser Lys Thr Glu Pro Gly Pro

165

170

175

Ala Ala Asp Ser Gln Lys Leu Ser Met Gly Gly Leu His Ser Ser Arg

180

185

190

Pro Leu Lys Arg Gly Leu Ser His Ser Leu Thr His Leu Asn Val Pro

195

200

205

Ser Thr Gly His Pro Ala Thr Ser Ala Pro His Thr Asn Gly Pro Gln
 210 215 220

Asp Leu Arg Pro Ser Thr Ser Gly Val Thr Phe Arg Ser Pro Leu Val
 225 230 235 240

Thr Ser Arg Ala Arg Ser Val Ser Ile Ser Val Pro Ser Thr Pro Arg
 245 250 255

Arg Gly Gly Ala Thr Gln Lys Pro Lys Pro Pro Trp Lys
 260 265

<210> 5

<211> 4248

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (183)..(2417)

<400> 5

tctcaaggtg catcactgtg catgggacaa atggcttgct aataaaagac accattgggt 60

ttgacacact aggtcattgt ttcttttttg aagatgggtat tgaacagagg aatactttgt 120

tccacaatct gggactcctc accaagccgg gtactctcct gccacccgat aggaacaact 180

cc atg tgt acc acc atg cga gat aaa gtg ttt gga aat tac att cct 227

Met Cys Thr Thr Met Arg Asp Lys Val Phe Gly Asn Tyr Ile Pro

1 5 10 15

gtg cct gct act gac tgt atg gct gtt tca act ttc tgg att gct cat 275

Val Pro Ala Thr Asp Cys Met Ala Val Ser Thr Phe Trp Ile Ala His

20 25 30

ccc aac aat aat ctg att aat aat gca gct gca ggc tca cag gat gct 323

Pro Asn Asn Asn Leu Ile Asn Asn Ala Ala Ala Gly Ser Gln Asp Ala

35 40 45

gga ata tgg tat tta ttc cac aag gaa cca act ggg gaa tcc agt gga 371

Gly Ile Trp Tyr Leu Phe His Lys Glu Pro Thr Gly Glu Ser Ser Gly

50 55 60

ttg cag ctc ttg gca aaa cca gaa ctc act cca ttg ggt ata ttt tat 419

Leu Gln Leu Leu Ala Lys Pro Glu Leu Thr Pro Leu Gly Ile Phe Tyr

65 70 75

aac aac agg gtc cat tca aat ttt aag gct ggc tta ttt att gac aaa 467

Asn Asn Arg Val His Ser Asn Phe Lys Ala Gly Leu Phe Ile Asp Lys

80 85 90 95

ggt gtc aaa aca acc aac tct agt gct gct gac cca agg gaa tac ctc 515

Gly Val Lys Thr Thr Asn Ser Ser Ala Ala Asp Pro Arg Glu Tyr Leu

100 105 110

tgt ttg gac aat agt gca aga ttt cga cct cat cag gat gca aac ccc 563

Cys Leu Asp Asn Ser Ala Arg Phe Arg Pro His Gln Asp Ala Asn Pro

115

120

125

gaa aaa cca cgt gtt gct gct cta att gac agg ctc att gct ttt aaa 611

Glu Lys Pro Arg Val Ala Ala Leu Ile Asp Arg Leu Ile Ala Phe Lys

130

135

140

aat aat gat aat gga gct tgg gtc aga gga gga gat att atc gtt caa 659

Asn Asn Asp Asn Gly Ala Trp Val Arg Gly Gly Asp Ile Ile Val Gln

145

150

155

aat tca gca ttt gca gat aat gga ata gga ctg acc ttt gcc agt gat 707

Asn Ser Ala Phe Ala Asp Asn Gly Ile Gly Leu Thr Phe Ala Ser Asp

160

165

170

175

gga agc ttc cca agt gat gaa ggt tcc agc caa gag gta tct gaa tct 755

Gly Ser Phe Pro Ser Asp Glu Gly Ser Ser Gln Glu Val Ser Glu Ser

180

185

190

ctc ttt gtt ggg gag agc agg aat tac ggc ttt cag ggt ggt cag aac 803

Leu Phe Val Gly Glu Ser Arg Asn Tyr Gly Phe Gln Gly Gly Gln Asn

195

200

205

aag tat gta ggc act gga gga ata gac cag aag cct cga aca tta ccc 851

Lys Tyr Val Gly Thr Gly Gly Ile Asp Gln Lys Pro Arg Thr Leu Pro

210

215

220

agg aac agg acg ttc cca att aga ggc ttt cag att tat gat ggg ccc 899

Arg Asn Arg Thr Phe Pro Ile Arg Gly Phe Gln Ile Tyr Asp Gly Pro

225

230

235

att cat ctc aca agg agc act ttc aaa aaa tat gtg cca act cca gat 947

Ile His Leu Thr Arg Ser Thr Phe Lys Lys Tyr Val Pro Thr Pro Asp

240

245

250

255

agg tac agc agt gca att ggc ttc ctc atg aag aat tcc tgg cag ata 995

Arg Tyr Ser Ser Ala Ile Gly Phe Leu Met Lys Asn Ser Trp Gln Ile

260

265

270

acc ccc agg aat aat atc tcc ctc gtg aag ttt ggt cca cat gtc tct 1043

Thr Pro Arg Asn Asn Ile Ser Leu Val Lys Phe Gly Pro His Val Ser

275

280

285

ctg aat gtc ttt ttt gga aag cct ggt ccc tgg ttt gaa gat tgt gag 1091

Leu Asn Val Phe Phe Gly Lys Pro Gly Pro Trp Phe Glu Asp Cys Glu

290

295

300

atg gat ggt gat aag aac tcc ata ttc cat gac att gat ggc tct gtg 1139

Met Asp Gly Asp Lys Asn Ser Ile Phe His Asp Ile Asp Gly Ser Val

305

310

315

aca gga tac aag gat gct tat gtg gga aga atg gac aac tac ctg atc 1187

Thr Gly Tyr Lys Asp Ala Tyr Val Gly Arg Met Asp Asn Tyr Leu Ile

320

325

330

335

cgc cat cca agc tgt gta aat gtg tct aag tgg aat gca gtg atc tgc 1235

Arg His Pro Ser Cys Val Asn Val Ser Lys Trp Asn Ala Val Ile Cys

340

345

350

agt ggg acc tat gca cag gtc tat gta cag aca tgg agc act cag aat 1283
 Ser Gly Thr Tyr Ala Gln Val Tyr Val Gln Thr Trp Ser Thr Gln Asn
 355 360 365

ctt tct atg acc att aca cga gat gag tat ccg tcc aac cct atg gtg 1331
 Leu Ser Met Thr Ile Thr Arg Asp Glu Tyr Pro Ser Asn Pro Met Val
 370 375 380

ctc cga ggt att aat cag aag gct gcc ttt cca cag tac cag cct gtc 1379
 Leu Arg Gly Ile Asn Gln Lys Ala Ala Phe Pro Gln Tyr Gln Pro Val
 385 390 395

gtc atg ctg gag aag ggt tat acc atc cac tgg aat ggg ccg gca cca 1427
 Val Met Leu Glu Lys Gly Tyr Thr Ile His Trp Asn Gly Pro Ala Pro
 400 405 410 415

cgg act aca ttt cta tac ctc gtc aac ttc aac aag aat gac tgg att 1475
 Arg Thr Thr Phe Leu Tyr Leu Val Asn Phe Asn Lys Asn Asp Trp Ile
 420 425 430

cga gtt ggc ctt tgc tat cca tca aac aca agt ttt caa gtt acc ttt 1523
 Arg Val Gly Leu Cys Tyr Pro Ser Asn Thr Ser Phe Gln Val Thr Phe
 435 440 445

ggc tat ttg cag cgg cag aat ggc tca tta tcc aaa atc gaa gaa tat 1571
 Gly Tyr Leu Gln Arg Gln Asn Gly Ser Leu Ser Lys Ile Glu Glu Tyr
 450 455 460

gag cct gtg cat tca ctg gaa gaa ctg caa aga aag caa tcc gag agg 1619

Glu Pro Val His Ser Leu Glu Glu Leu Gln Arg Lys Gln Ser Glu Arg

465

470

475

aaa ttc tat ttt gac tcc agc acg ggg tta ctg ttt ttg tat ctc aaa 1667

Lys Phe Tyr Phe Asp Ser Ser Thr Gly Leu Leu Phe Leu Tyr Leu Lys

480

485

490

495

gcc aaa agc cac agg cat ggc cac agt tac tgt tca tct cag gga tgt 1715

Ala Lys Ser His Arg His Gly His Ser Tyr Cys Ser Ser Gln Gly Cys

500

505

510

gaa aga gtc aag atc caa gca gcc aca gac tca aag gac atc agt aac 1763

Glu Arg Val Lys Ile Gln Ala Ala Thr Asp Ser Lys Asp Ile Ser Asn

515

520

525

tgc atg gcc aaa gca tac cca cag tac tac aga aag ccg tca gtg gtc 1811

Cys Met Ala Lys Ala Tyr Pro Gln Tyr Tyr Arg Lys Pro Ser Val Val

530

535

540

aag cgg atg ccg gcc atg ctc act gga ctc tgt caa ggc tgt ggc act 1859

Lys Arg Met Pro Ala Met Leu Thr Gly Leu Cys Gln Gly Cys Gly Thr

545

550

555

cgg cag gtg gtg ttt act agt gat cct cat aaa agt tac ctc cct gtg 1907

Arg Gln Val Val Phe Thr Ser Asp Pro His Lys Ser Tyr Leu Pro Val

560

565

570

575

caa ttc cag tca cct gat aaa gca gaa gcc cag cgt gga gac ccg tct 1955

Gln Phe Gln Ser Pro Asp Lys Ala Glu Ala Gln Arg Gly Asp Pro Ser

580

585

590

gtt att tct gtc aat ggc act gac ttt acc ttc cga agt gca ggc gtc 2003

Val Ile Ser Val Asn Gly Thr Asp Phe Thr Phe Arg Ser Ala Gly Val

595

600

605

ctc ctc ctt gtt gtg gat ccg tgc agc gtt cca ttc cgc ttg acg gaa 2051

Leu Leu Leu Val Val Asp Pro Cys Ser Val Pro Phe Arg Leu Thr Glu

610

615

620

aaa acg gtt ttt cct ctt gct gat gtc agt cgc att gaa gag tat tta 2099

Lys Thr Val Phe Pro Leu Ala Asp Val Ser Arg Ile Glu Glu Tyr Leu

625

630

635

aaa aca ggc atc cct cca agg tcc att gtt ctg ttg agc aca aga gga 2147

Lys Thr Gly Ile Pro Pro Arg Ser Ile Val Leu Leu Ser Thr Arg Gly

640

645

650

655

gaa ata aag cag tta aac att tca cac tta cta gta cct ctg gga tta 2195

Glu Ile Lys Gln Leu Asn Ile Ser His Leu Leu Val Pro Leu Gly Leu

660

665

670

gcc aaa cca gct cat ctt tat gac aaa ggg agt acc gta ttt ttg gga 2243

Ala Lys Pro Ala His Leu Tyr Asp Lys Gly Ser Thr Val Phe Leu Gly

675

680

685

ttc agt gga aac ttt aaa cca tca tgg act aag cta ttt acc agt cct 2291

Phe Ser Gly Asn Phe Lys Pro Ser Trp Thr Lys Leu Phe Thr Ser Pro

690

695

700

gct gga cag ggc ctt ggg gtg ctt gaa caa ttc ata cct ttg cag ctg 2339

Ala Gly Gln Gly Leu Gly Val Leu Glu Gln Phe Ile Pro Leu Gln Leu

705

710

715

gac gaa tat ggt tgt ccc aga gcc acc act gtc cgc aga aga gac ctg 2387

Asp Glu Tyr Gly Cys Pro Arg Ala Thr Thr Val Arg Arg Arg Asp Leu

720

725

730

735

gaa ctg cta aag caa gct tca aaa gca cat tagagactaa ctgtaactta 2437

Glu Leu Leu Lys Gln Ala Ser Lys Ala His

740

745

agtgctgggg gaaaaaaaat gtgaactaac ttatttaatt tatggcattt taaaatgaca 2497

ctgttaaccc aacggaacca ttttccagtt tgatacagaa tggggagaaa agaaagcggt 2557

tgaaattatt gcttgatac cagcttcatg caccttctag ttgtacaaaa tgttaaagac 2617

gttgtttgta ttgttaaggc tgggtgtattc agagagcaga tctcttattc ctcactttcc 2677

acccccgtat ttgttaatga ccatgagcaa tgtttttact ttttgtataa tgggggtgggg 2737

tggagtgggg gcttctgaga gtcagcctga ggtctttaga ggaccagcta ttgtagcacc 2797

ttggatactt gaagttaaat gctcagttgg gtcgggtggc agttgacttg gtggctggca 2857

tgttcagcag tgcctggggc cctgtttctg ggcagccttt gaggattttc tatgatattg 2917

aatgacagtt ttaagtggca actcaggccc agctcatgcc cttttttgcc tggacatgtg 2977

ctatttttat tcacttatat atcaattact tgtaagggtt aaactttcaa acaggaagta 3037

tattgggaca aaagggcctt tggggattag atatcccttt aatctgtgac cattgggcaa 3097

aaaattttcc tgcagcaaaa gtctgaggct gttagggacca tttttgcagc tttaatcctt 3157

agcctctttt gactgtatat ttgtgtttaa aatgcagagc tcaactgaat atttcctttt 3217

ttgttttgtt ttgttttgtt ttaagaagta ggttgttttc ctgaaccgta aacttgtatc 3277

attttaactt gcacaaagga agtctgttct tggtattgct cttgcacttg ggttttttgt 3337

tattgttttg tgtggatttt ttaaagcitt tctgttcacc ctctgccag gaaaatccca 3397

gaaagcttaa tgatacccca aaatgattac acccaggag gaaaaaaagg agcgctttct 3457

agggtcagaa tcgtggagag aatactcaga aatgaacctc tttaaagcct tgcaggaatg 3517

agtcactctt acttaatgaa atgttaaagc caattaaaaa gcatgctgtg atgccagct 3577

tccctttcca cagggtgcat gcgtctcctg ctggtgaatc acatgcggca agaggcaact 3637

ggctccacag cctgggatgc tgccgtacca agaggaaaga agcagcaaaa tgccttttacg 3697

ttgttctaaa cccccgacgc ataaagtgtg gaggagggat ggccaagggt gggtggtaga 3757

aagtgtgttc aggctgacac tggcaatgag tacagataat ttcactttcc tcttctaggg 3817

gcaaaggctg atggcctcta cctttgtatc caggagaaac tgcagagcag ccctgtgact 3877

ttacaaaata tgctacctca aagtgtacc gataaacctt tctaattgta agtgcctta 3937

ctaagggcac atgtcttaat caaagttagt tttttgtttt ctggtttgtt ttttttttt 3997

gtatattgat gaatgagatc ttacctatta aatatgttat tggattatgg ttcctgaagg 4057

tcattagagt gtgtgtgtgt gtgtgtgtgt gtgtgtgttt tatgacttaa atatctttac 4117

gtgtgttttt tagagcttgg ttcttttaaag atttgagaa gatatgtaaa ttaccaaggc 4177

acttggtttt tctgttttat atactaataa tcagggccta agttaataa aaatatgtgt 4237

gcatgtattt t 4248

<210> 6

<211> 745

<212> PRT

<213> Homo sapiens

<400> 6

Met Cys Thr Thr Met Arg Asp Lys Val Phe Gly Asn Tyr Ile Pro Val

1

5

10

15

Pro Ala Thr Asp Cys Met Ala Val Ser Thr Phe Trp Ile Ala His Pro

20

25

30

Asn Asn Asn Leu Ile Asn Asn Ala Ala Ala Gly Ser Gln Asp Ala Gly

35

40

45

Ile Trp Tyr Leu Phe His Lys Glu Pro Thr Gly Glu Ser Ser Gly Leu

50

55

60

Gln Leu Leu Ala Lys Pro Glu Leu Thr Pro Leu Gly Ile Phe Tyr Asn

65

70

75

80

Asn Arg Val His Ser Asn Phe Lys Ala Gly Leu Phe Ile Asp Lys Gly

85

90

95

Val Lys Thr Thr Asn Ser Ser Ala Ala Asp Pro Arg Glu Tyr Leu Cys

100

105

110

Leu Asp Asn Ser Ala Arg Phe Arg Pro His Gln Asp Ala Asn Pro Glu

115

120

125

Lys Pro Arg Val Ala Ala Leu Ile Asp Arg Leu Ile Ala Phe Lys Asn

130

135

140

Asn Asp Asn Gly Ala Trp Val Arg Gly Gly Asp Ile Ile Val Gln Asn

145

150

155

160

Ser Ala Phe Ala Asp Asn Gly Ile Gly Leu Thr Phe Ala Ser Asp Gly

165

170

175

Ser Phe Pro Ser Asp Glu Gly Ser Ser Gln Glu Val Ser Glu Ser Leu
180 185 190

Phe Val Gly Glu Ser Arg Asn Tyr Gly Phe Gln Gly Gly Gln Asn Lys
195 200 205

Tyr Val Gly Thr Gly Gly Ile Asp Gln Lys Pro Arg Thr Leu Pro Arg
210 215 220

Asn Arg Thr Phe Pro Ile Arg Gly Phe Gln Ile Tyr Asp Gly Pro Ile
225 230 235 240

His Leu Thr Arg Ser Thr Phe Lys Lys Tyr Val Pro Thr Pro Asp Arg
245 250 255

Tyr Ser Ser Ala Ile Gly Phe Leu Met Lys Asn Ser Trp Gln Ile Thr
260 265 270

Pro Arg Asn Asn Ile Ser Leu Val Lys Phe Gly Pro His Val Ser Leu
275 280 285

Asn Val Phe Phe Gly Lys Pro Gly Pro Trp Phe Glu Asp Cys Glu Met
290 295 300

Asp Gly Asp Lys Asn Ser Ile Phe His Asp Ile Asp Gly Ser Val Thr
305 310 315 320

Gly Tyr Lys Asp Ala Tyr Val Gly Arg Met Asp Asn Tyr Leu Ile Arg
325 330 335

His Pro Ser Cys Val Asn Val Ser Lys Trp Asn Ala Val Ile Cys Ser

340

345

350

Gly Thr Tyr Ala Gln Val Tyr Val Gln Thr Trp Ser Thr Gln Asn Leu

355

360

365

Ser Met Thr Ile Thr Arg Asp Glu Tyr Pro Ser Asn Pro Met Val Leu

370

375

380

Arg Gly Ile Asn Gln Lys Ala Ala Phe Pro Gln Tyr Gln Pro Val Val

385

390

395

400

Met Leu Glu Lys Gly Tyr Thr Ile His Trp Asn Gly Pro Ala Pro Arg

405

410

415

Thr Thr Phe Leu Tyr Leu Val Asn Phe Asn Lys Asn Asp Trp Ile Arg

420

425

430

Val Gly Leu Cys Tyr Pro Ser Asn Thr Ser Phe Gln Val Thr Phe Gly

435

440

445

Tyr Leu Gln Arg Gln Asn Gly Ser Leu Ser Lys Ile Glu Glu Tyr Glu

450

455

460

Pro Val His Ser Leu Glu Glu Leu Gln Arg Lys Gln Ser Glu Arg Lys

465

470

475

480

Phe Tyr Phe Asp Ser Ser Thr Gly Leu Leu Phe Leu Tyr Leu Lys Ala

485

490

495

Lys Ser His Arg His Gly His Ser Tyr Cys Ser Ser Gln Gly Cys Glu

500

505

510

Arg Val Lys Ile Gln Ala Ala Thr Asp Ser Lys Asp Ile Ser Asn Cys

515

520

525

Met Ala Lys Ala Tyr Pro Gln Tyr Tyr Arg Lys Pro Ser Val Val Lys

530

535

540

Arg Met Pro Ala Met Leu Thr Gly Leu Cys Gln Gly Cys Gly Thr Arg

545

550

555

560

Gln Val Val Phe Thr Ser Asp Pro His Lys Ser Tyr Leu Pro Val Gln

565

570

575

Phe Gln Ser Pro Asp Lys Ala Glu Ala Gln Arg Gly Asp Pro Ser Val

580

585

590

Ile Ser Val Asn Gly Thr Asp Phe Thr Phe Arg Ser Ala Gly Val Leu

595

600

605

Leu Leu Val Val Asp Pro Cys Ser Val Pro Phe Arg Leu Thr Glu Lys

610

615

620

Thr Val Phe Pro Leu Ala Asp Val Ser Arg Ile Glu Glu Tyr Leu Lys

625

630

635

640

Thr Gly Ile Pro Pro Arg Ser Ile Val Leu Leu Ser Thr Arg Gly Glu

645

650

655

Ile Lys Gln Leu Asn Ile Ser His Leu Leu Val Pro Leu Gly Leu Ala

660

665

670

Lys Pro Ala His Leu Tyr Asp Lys Gly Ser Thr Val Phe Leu Gly Phe

675

680

685

Ser Gly Asn Phe Lys Pro Ser Trp Thr Lys Leu Phe Thr Ser Pro Ala

690

695

700

Gly Gln Gly Leu Gly Val Leu Glu Gln Phe Ile Pro Leu Gln Leu Asp

705

710

715

720

Glu Tyr Gly Cys Pro Arg Ala Thr Thr Val Arg Arg Arg Asp Leu Glu

725

730

735

Leu Leu Lys Gln Ala Ser Lys Ala His

740

745

<210> 7

<211> 1545

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (341)..(1006)

<400> 7

tcattttttt tctttgctta atttgaccaa gaataagttc attcccaaac tctcacctta 60

tgcttagctc tcctctcaat attcaaactc ctagaacacct gttgttttgt ctttcccatg 120

gacacggctc ccaggcctct gacccctgct ctaattggga cctgctgtgt ggccctccct 180

tgcttaccag ctgacaggaa ccttictca cccccagggt ggacacgccg tttccaaggc 240

ctcatggctt ctttttctt ggttactgcc tcgggctccc tgggagagat ctttttggtg 300

ccgaaaaccg gaacgggaag cctcagcacc ctggcccccc atg ccc ctc gtg ggg 355

Met Pro Leu Val Gly

1 5

cag ggt ggg tat acc ctg tac acc ctc ctg gtt tgg gct gag ggc att 403

Gln Gly Gly Tyr Thr Leu Tyr Thr Leu Leu Val Trp Ala Glu Gly Ile

10 15 20

agg ggt acc ggg cgt ctt tgg ggt ggc att agc ccc cga gtt gct ttg 451

Arg Gly Thr Gly Arg Leu Trp Gly Gly Ile Ser Pro Arg Val Ala Leu

25 30 35

gaa act gtc ata ctt tct tct gtc tta gaa ctc aga atc caa gaa atg 499

Glu Thr Val Ile Leu Ser Ser Val Leu Glu Leu Arg Ile Gln Glu Met

40 45 50

gca tcc atg ggg ata ggc aac cag cca ttc atg gat gtc aag ccc aga 547

Ala Ser Met Gly Ile Gly Asn Gln Pro Phe Met Asp Val Lys Pro Arg

55

60

65

gac cgg acc cct gac tgt gca gtg ata agc gac ggg gct ccc aaa tgt 595

Asp Arg Thr Pro Asp Cys Ala Val Ile Ser Asp Gly Ala Pro Lys Cys

70

75

80

85

gca gtg atg agc gac cgg gtc ccc gac agc atc gtc aag ggc aca ggt 643

Ala Val Met Ser Asp Arg Val Pro Asp Ser Ile Val Lys Gly Thr Gly

90

95

100

acg gtg gct cgg tcc cgc cct cac tca ccc tgc aga ggg cac tgg gcc 691

Thr Val Ala Arg Ser Arg Pro His Ser Pro Cys Arg Gly His Trp Ala

105

110

115

tgt cat caa ggg cat ggg tac ggc ggc atc ggc ccc acc ctc act cgc 739

Cys His Gln Gly His Gly Tyr Gly Gly Ile Gly Pro Thr Leu Thr Arg

120

125

130

cca cag agt gca cca ggc ctg tcg tca agg aca cgg gta cgg tgg cct 787

Pro Gln Ser Ala Pro Gly Leu Ser Ser Arg Thr Arg Val Arg Trp Pro

135

140

145

cgg ccc cgc cct cac tca tcc tgc aga ggg cac tgg gcc agt ggc cga 835

Arg Pro Arg Pro His Ser Ser Cys Arg Gly His Trp Ala Ser Gly Arg

150

155

160

165

cat ggt ggg ttg gat ggg cat gac tgc agt ggc aaa gcc tgg tcg gcc 883

His Gly Gly Leu Asp Gly His Asp Cys Ser Gly Lys Ala Trp Ser Ala

170

175

180

ttt cag acg gct ctg atc cca ttc ccg aac ctg ggc tgc act tca gga 931

Phe Gln Thr Ala Leu Ile Pro Phe Pro Asn Leu Gly Cys Thr Ser Gly

185

190

195

gcg gaa gcc agc ctg acg tgc ttt ctg tcc ctt tcc aga gtc aca aat 979

Ala Glu Ala Ser Leu Thr Cys Phe Leu Ser Leu Ser Arg Val Thr Asn

200

205

210

gag agg gtc cac agc ggt gtc ctc ctc tgaccagcc gcccccttca 1026

Glu Arg Val His Ser Gly Val Leu Leu

215

220

agcgaccaca ctccaccatc tcagacagca gcacctctc ttctagcagc cagtcctcct 1086

ccatcctggg gtcgctgggc ctgcttgtgt cctccagccc agcccacccg ggcctatgga 1146

gccctgccc aagcccctgg tcatttgata tcagagcctg cgtcgaggaa gatgagccag 1206

agccagaact agagacgggc acccaggctg cagtgtgtga gggggctcct gctgtgctgc 1266

tgagccgcac acgccaggcc tgatgactgt cagggtggca gtgcccatca tgtggctaga 1326

acaatacaga gggagcagca cgccaggcct gatgactctg ggggtggcgg tgcccatcgc 1386

gtggctggaa cgatccagag ggaatagcac agcagggtgc caggattttt 1446

agacattccc taacattttc aaacaaattt acaattttgt cttatttaaa aaacaaacct 1506

tccacttcca cccaagacaa cagcatagga aacagacct 1545

<210> 8

<211> 222

<212> PRT

<213> Homo sapiens

<400> 8

Met Pro Leu Val Gly Gln Gly Gly Tyr Thr Leu Tyr Thr Leu Leu Val

1 5 10 15

Trp Ala Glu Gly Ile Arg Gly Thr Gly Arg Leu Trp Gly Gly Ile Ser

20 25 30

Pro Arg Val Ala Leu Glu Thr Val Ile Leu Ser Ser Val Leu Glu Leu

35 40 45

Arg Ile Gln Glu Met Ala Ser Met Gly Ile Gly Asn Gln Pro Phe Met

50 55 60

Asp Val Lys Pro Arg Asp Arg Thr Pro Asp Cys Ala Val Ile Ser Asp

65 70 75 80

Gly Ala Pro Lys Cys Ala Val Met Ser Asp Arg Val Pro Asp Ser Ile

85 90 95

Val Lys Gly Thr Gly Thr Val Ala Arg Ser Arg Pro His Ser Pro Cys

100

105

110

Arg Gly His Trp Ala Cys His Gln Gly His Gly Tyr Gly Gly Ile Gly

115

120

125

Pro Thr Leu Thr Arg Pro Gln Ser Ala Pro Gly Leu Ser Ser Arg Thr

130

135

140

Arg Val Arg Trp Pro Arg Pro Arg Pro His Ser Ser Cys Arg Gly His

145

150

155

160

Trp Ala Ser Gly Arg His Gly Gly Leu Asp Gly His Asp Cys Ser Gly

165

170

175

Lys Ala Trp Ser Ala Phe Gln Thr Ala Leu Ile Pro Phe Pro Asn Leu

180

185

190

Gly Cys Thr Ser Gly Ala Glu Ala Ser Leu Thr Cys Phe Leu Ser Leu

195

200

205

Ser Arg Val Thr Asn Glu Arg Val His Ser Gly Val Leu Leu

210

215

220

<210> 9

<211> 30

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:an artificially
synthesized oligo-cap linker sequence

<400> 9

agcaucgagu cggccuuguu ggccuacugg

30

<210> 10

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:an artificially
synthesized oligo(dT) primer sequence

<400> 10

gcggctgaag acggcctatg tggccttttt tttttttttt tt

42

<210> 11

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:an artificially

synthesized primer sequence

<400> 11

agcatcgagt cggccttggt g

21

<210> 12

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:an artificially
synthesized primer sequence

<400> 12

gcggctgaag acggcctatg t

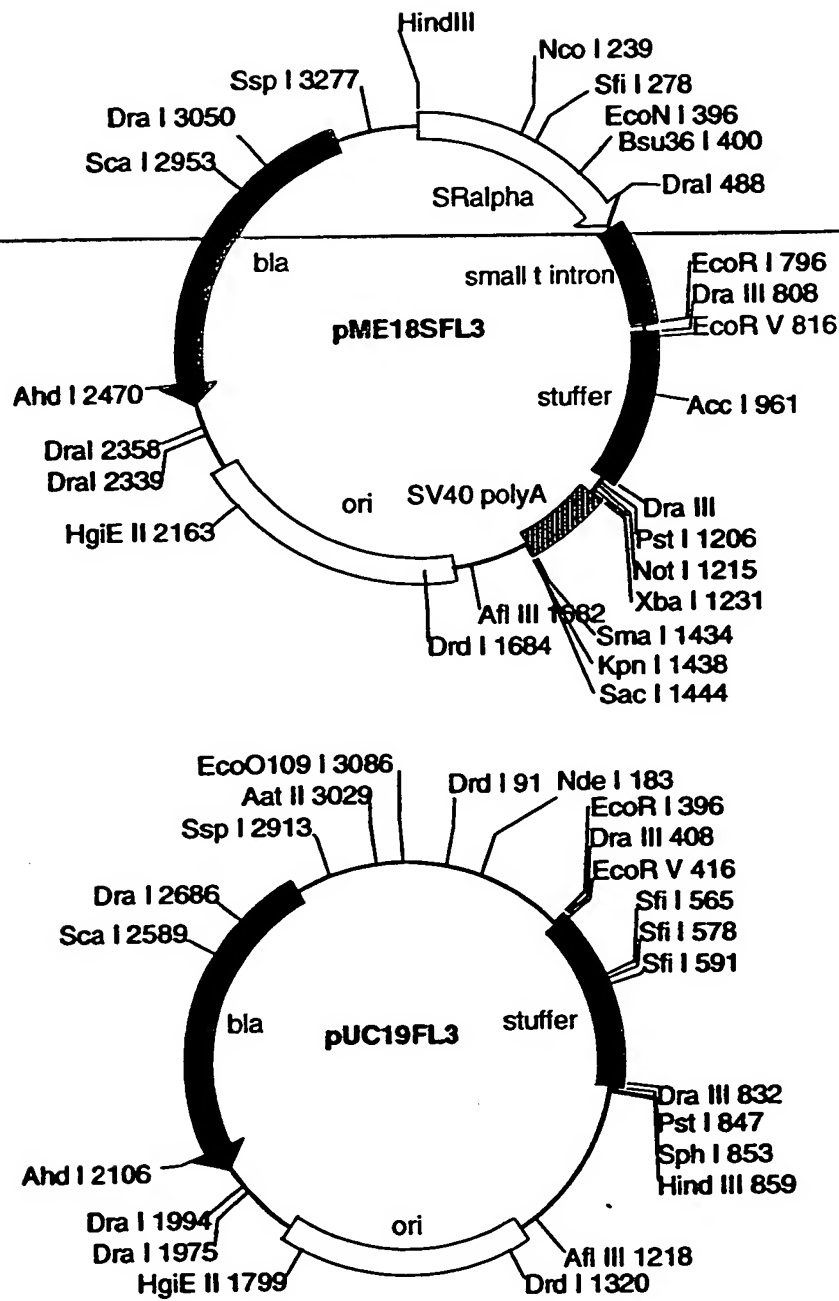
21

【図面の簡単な説明】

【図 1】 pME18SFL3とpUC19FL3のベクターのマップ

【書類名】 図面

【図 1】



【書類名】 要約書

【要約】

【課題】 新規なヒト・タンパク質と、それをコードする全長cDNAの提供。

【解決手段】 4種のヒト・タンパク質とそれをコードする全長cDNAを単離した。
。また本発明による全長cDNAは、これらのタンパク質の生産のために利用される

【選択図】 なし

出 願 人 履 歴 情 報

識別番号

[597059742]

1. 変更年月日	1997年 4月28日
[変更理由]	新規登録
住 所	千葉県木更津市矢那1532番地3
氏 名	株式会社ヘリックス研究所